

DOCUMENT RESUME

ED 243 944

TM 840 257

AUTHOR Lapointe, Archie E.
TITLE Danger: Work on Higher Levels.
PUB DATE 25 Apr 84
NOTE 13p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New Orleans, LA, April 24-26, 1984).
PUB TYPE Speeches/Conference Papers (150) -- Viewpoints (120)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Abstract Reasoning; Computer Literacy; Computer Science; *Educational Assessment; *Educational Objectives; Evaluative Thinking; *Federal Programs; Quantitative Tests; Reading Tests; Science Tests; Synthesis; *Test Construction
IDENTIFIERS *Blooms Taxonomy; *National Assessment of Educational Progress

ABSTRACT

The Assessment Policy Committee of the National Assessment of Educational Progress (NAEP) has directed the NAEP staff to focus the 1985-86 Assessments of Reading, Mathematics, Science, and Computer Competence on the higher-order skills. Each Learning Area Committee worked independently developing three-dimensional models. These defined what could be known (a single-step cognitive process), what could be used (two cognitive steps), and what could be synthesized and applied to new situations (three or more steps). While it was decided to exploit the discoveries of modern cognitive theory, every attempt will be made to define categories so that references back to the Bloom taxonomy used in the last two assessments will be possible. The statements of learning objectives that were hammered out by each Committee are currently being reviewed by scores of experts, teachers, and lay people. When this step is complete, each Learning Area Committee will be reconvened to finalize the objectives. (BW)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED243944

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Archie E. Lapointe

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

"Danger: Work on Higher Levels"*

As a boy, I listened faithfully to the radio adventures of "The Shadow" every Sunday afternoon. Each week, the announcer advised us that Lamont Cranston had developed his powers to "cloud men's minds" during a trip to the mysterious East.

When the Assessment Policy Committee of the National Assessment of Educational Progress directed NAEP's staff to focus the 1985-86 Assessments of Reading, Mathematics, Science and Computer Competence on the higher-order skills, I too decided to go searching for wisdom in the Orient. My wife and I sallied forth to Hong Kong and to mainland China in search of Truth. It was a marvelous trip and we had a delightful time. As we wandered through the teeming labyrinthine streets and alleys of the thriving and still growing Royal Colony, I kept all of my senses alert to clues that would help me with my dilemma. Alas, all I found were countless warnings, ubiquitously displayed as though someone had been sent on ahead of me to counsel me to give up the search. Or else, they, the Chinese, have been trying to solve the puzzle for 6,000 years and are charitably discouraging others from the effort.

In any event, they are not subtle. Dozens of new buildings are under construction throughout Hong Kong and on the scaffolding surrounding these sites, are countless red and white signs with the clear warning in Chinese and in English: "Danger: Work on Higher Levels."

*AERA Presentation, April 25, 1984, New Orleans, Louisiana

Discouraged, but not dismayed, we returned to Princeton, searched the literature, talked to experts and called together some committee members to meditate on the riddle:

"How can honorable assessors measure that which has no definition?"

Buddha was out so we studied Bloom and considered Cronbach instead of Confucius. Monks were replaced by mathematicians, saints by statisticians and during a weekend in January, we ascended the mountain.

But, let me give you some background.

During the past 10 years, NAEP data have fairly consistently shown that while achievement scores on those exercises that measure basic reading, math, writing and science skills have been improving at the three age levels assessed by NAEP (9-13-17), the same assessments have recorded consistent declines in students' abilities to infer significance from a printed passage, to classify and solve math problems, to marshal arguments in support of a thesis or to interpret the possible meanings of scientific data. These findings are consistent with results noted by analyzing SAT and ACT results as well as some widely used commercial tests.

- o The "basics" are under control. Most students at all three age levels have a good command of comprehension skills, math basic operations and writing mechanics. This is true in absolute terms, i.e., they select correct answers to questions

measuring these skills and there has been significant improvement over the past 10 years in performance on the same exercises.

- o There has been consistent decline in the achievement levels of the more able students (i.e., those who score in the top quartile on the assessment itself) in reading, math and science reflecting weaker performance on items that attempt to measure inference, problem-solving and analytical abilities
- o A single specific example of the facts is the result of the most recent math assessment in 1982. These data are illustrated on chart A which shows growth for Black, low-achieving students at all three age levels and declines for high-achieving, majority students at all three age levels.
- o In writing, students' ability to support a position with clear and relevant arguments is weak at all age levels.
- o Few students at each age level are able to grasp the significance of the tone or mood of a piece of good writing or infer significance much beyond the articulated points in a passage.
- o In science, 9-year-olds scored higher on items dealing with conclusions based on observation than on items requiring the more complex processes of interpolation and extrapolation.

Among 17-year-olds 83% were able to classify geometric objects by a single dimension, but only 52% could classify them using two dimensions.

The good news, of course, is that we seem to have demonstrated that we are in control of the situation. We have been stressing "the basics" during the past decade and we have been successful. We have neglected developing the potential of the gifted and talented, and NAEP and SAT results suggest that their performance on exercises that attempt to measure higher-order skills has declined consistently during the period 1970-1980. Our directed energies seem to have produced success and our neglect parallels failure.

Observation of the repetition of the phenomenon across curriculum areas suggests a commonality or sameness of skill. The processes involved in deciphering the meaning of a complex passage or poem may be the same as those employed to break down the elements of a math problem or to develop a scientific hypothesis from a set of data. We think NAEP's new design will permit us to learn something about this possibility as well.

First, let's challenge some basic assumptions. What's a nice project like NAEP doing even pursuing the question? Isn't NAEP a survey program designed to merely describe the student population at three age levels and report on trends over time? That is certainly an essential part of NAEP's mandate. However, an important element of NAEP's design from the outset - clearly in the minds of Ralph Tyler, Frank Keppel and John Tukey - was the notion of using NAEP to advance the technology of measurement. The Assessment's objective-referenced exercises, matrix sampling and one-on-one administration of items were exciting innovations 10 years ago.

If current policy and decision makers need to know more about students' abilities with higher-order skills, it seems in keeping with NAEP's tradition to attempt to contribute to addressing the issue. Furthermore, new design features recently implemented make the assessment an ideal vehicle for comparing skills across curriculum areas.

But back to my story.

The charge having been issued by the APC to focus on these skills, Ina Mullis and her staff developed several ideas for addressing them and we convened the four Learning Area Committees (Reading, Math, Science and Computer Competence) in Princeton during the same three day period in the same farm house. NAEP committees had never met together before. On day one, all members of the four groups were brought together in the same room and given the same charge, viz. to focus on this higher end in designing their particular assessments and to develop knowledge and skill matrices for each subject that might permit developing exercises across areas.

Each committee then worked independently developing what ended up being three-dimensional models or cubes rather than matrices. These defined what could be known (a single-step cognitive process), what could be used (two cognitive steps) and what could be synthesized and applied to new situations (three or more steps). While Bloom's taxonomy served as an initial discussion springboard, current research into theories of cognition married to logic and to the awareness of research, and solid classroom experience of the Committee members shaped the decisions of each curriculum group.

Near the end of the three days, all were reconvened to present their respective results to the assembly. There was genuine excitement when they observed the obvious parallel patterns that had evolved.

There was little, if any challenge to the priority being given this higher end of the ability spectrum. On the contrary, there was easy concurrence that relatively little value would be derived from simply discovering and reporting how much students knew or how well they could perform basic functions. There was quick agreement that the needs in today's and tomorrow's environments will be to classify, analyse, synthesize and apply knowledge. Loading a human or electronic memory with information is relatively straightforward. Helping young people develop techniques for scanning long-term memory, for classifying and sorting new information, for developing and testing hypotheses and for applying principles are the challenges of today's instructional programs.

While it was decided to exploit the discoveries of modern cognitive theory, every attempt will be made to define categories so that references back to the Bloom taxonomy used in the last two assessments will be possible.

Here are some specifics about the goals each of the curriculum teams have set for themselves. Keep in mind that each student being assessed will be asked to respond to questions in more than one subject so it will be possible to observe relationships across curricula.

The mathematics group felt strongly that while NAEP should play some kind of leadership role, its main task was describing reality and recognizing achievement and success. Teachers need the recognition, encouragement and feedback that they are doing a good job and that their activities bear fruit. Important topics such as knowledge of numbers and operations and ability to perform basic skills continue to be important though they are less attractive than others in today's media-hype discussions. These will be assessed.

Stress will, nevertheless, probably be on the content areas such as Geometry, Measurement, Data Organization and Interpretation, Discrete Mathematics and Mathematical Methods. The cognitive levels that will be probed include Understanding and Comprehension, Routine Applications and Problem Solving/Reasoning. Relative focus on certain of these aspects will depend upon age and grade level and will reflect what is known about typical development stages as well as probable exposure to content.

The science team defined its world using three broad terms: Content, Context and Cognition. They hope to measure the first within the second, while aware of the third. To be specific, examples of Content areas include Life Sciences, Physics, Chemistry, Earth and Space Sciences, History of Science and Nature of Science. The Contexts in which each of these bodies of knowledge are typically applied are Personal, Societal, Technological and Scientific. Certain students will be able to demonstrate competencies in all of these contexts. A given student can be effective in each of these pairs of content and context at various Cognitive levels: knowledge, ability to use knowledge or ultimately to be able to synthesize, integrate and apply knowledge. An attempt will be made to plumb each of these possibilities.

Our reading experts warmed to the task by listing a progressive set of cognitive steps. For each, they described some sample tasks that could possibly measure each level. For example:

<u>Level</u>	<u>Sample Tasks</u>
Identify Information	To locate To confirm To affirm
Integrate/Use Information	To restructure To compare To reformulate
Extend Information	To generalize To conclude To predict To interpret
Make Critical Judgments	To evaluate To react To use literary conventions

Finally the task of the computer competence committee was both more difficult because it represents an uncharted area in which all kinds of programs are evolving and easier because there is no restriction imposed by an existing data base. They were unanimous in their intention to assess Computer Competence rather than Computer Understanding because the latter name implies

a secondhand rather than a direct knowledge of computers. They organized the corpus of their set of experiences into Content which includes such items as:

- o Programming Language Elements
- o Data Constants, Variables, Structures
- o Control Structures, Loops, Branches, Subroutines
Procedures
- o Principles of Program Planning, Design and Testing
- o Study of Algorithms

and Cognitive levels which they identify using such vocabulary as:

- o Knowledge
- o Skills
- o Understanding/Comprehension/Analysis
- o Application/Synthesis
- o Problem Solving/Reasoning

What are the next steps?

The real challenge, of course, lies ahead. While NAEP is not in the business of setting standards, the Federal legislation that created NAEP, mandates that learning objectives be established through a consensus process. We take that charge seriously. The statements of learning objectives that were hammered out by each Committee are currently being reviewed by scores of experts, teachers and lay people.

When this step is complete, each Learning Area Committee will be reconvened to finalize the objectives.

Meanwhile, the literature is being searched for models that could be helpful. Experts are being consulted for suggestions about how we can effectively measure these skills, lists of clever item writers are being drafted and techniques for creating exercises in brain storming settings are being planned. We have invited Clare Burstall's colleagues to discuss their experiences with us and they have generously shared their time, materials and candidly expressed their conclusions.

Fortunately, NAEP is not limited to multiple choice questions. As a matter of fact, we are committed to using some open ended or essay questions in each curriculum area. If there is solid justification for one-on-one testing or small group assessment as used in the United Kingdom, funding will be sought to implement new techniques.

We fully recognize the set of issues that accompany new or different item types. There are questions of standardized administration and reliability of scoring. Since NAEP uses trained exercise administrators, the first is manageable and NAEP's long history of essay scoring enables us to anticipate the problems of creating scoring guides, training and managing the people who do it and maintaining reliability in the process.

The technical questions can be addressed. The real issues have to do with the imagination, creativity and sheer genius necessary to dream up the processes. Ongoing research is scattered and successful models are hard to find.

NAEP has immediate access to the work of Norman Fredericksen and Bill Ward at ETS as well as the guidance of Sam Messick and the accumulated test development experience from programs that range from the SAT and Advanced Placement, through the Graduate Record Exams. In-basket techniques as well as research done with the Navy and Medical schools will be available as we refine the assessment. Preliminary surveys of all of this talent and experience suggest a frightening conclusion; there's a lot of pioneering work to be done and field tests may yield disappointing results.

We are intent on giving it a solid effort. If we are successful and analyses of the data indicate validity and reportable conclusions, our hope is that you will all share the benefits since NAEP exercises and methodology are in the public domain. If, in the process of the effort, NAEP staff could advance the state of the measurement art a little bit, that would constitute a psychological perk that would make coming to work even more satisfying.

As a non-psychometrician, non-statistician, non-research scientist, my unearned reward has been, is and promises to continue to be, standing under the scaffolding while all of this magnificent construction is going on all around me -- completely oblivious to the clearly posted caveat:

Danger: Work on Higher Levels

Mathematics Performance

